

HELSINGIN YLIOPISTO

Työn hallittavuus-vaativuus -kyselyn rakennevaliditeetti ja mittausinvarianssi

Sakari Lintula
Pro gradu -tutkielma
Psykologia
Lääketieteellinen tiedekunta
Helsingin Yliopisto
Joulukuu 2019
Ohjaaja: Jari Lipsanen

Tiedekunta – Fakultet – Faculty Lääketieteellinen tiedekunta		Koulutusohjelma – Utbildningsprogram – Degree Programme Psykologia
Tekijä – Författare – Author Sakari Johannes Lintula		
Työn nimi – Arbetets titel – Title Työn hallittavuus-vaativuus -kyselyn rakennevaliditeetti ja mittausinvarianssi Suomessa		
Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track Psykologia		
Työn laji – Arbetets art – Level Pro gradu -tutkielma	Aika – Datum – Month and year Joulukuu 2019	Sivumäärä – Sidoantal – Number of pages 27 + 2 liitettä (2 ls)
<p>Tiivistelmä – Referat – Abstract</p> <p>Mikäli kykenemme tunnistamaan työstressiä edeltävät tekijät, voimme löytää keinoja vaikuttaa työssäkäyvien ihmisten mielenterveyteen ennakoivasti. Näiden tekijöiden tunnistamiseen vaaditaan validia mittaamista. Tässä tutkimuksessa keskitytään työn hallittavuus-vaativuus -kyselyn suomenkielisen patteriston validiteettiin, jota on tutkittu maailmalla ja Suomessa. Kuitenkin epäselvyyksiä on jäljellä ja niitä pyritään selvittämään tässä tutkimuksessa.</p> <p>Otoksena toimii Lasten Sepelvaltimotaudin Riskitekijät -tutkimuksessa (LASERI) kerätty aineisto, joka sisältää laaja-alaisesti eri ammattiryhmissä olevia henkilöitä. Aineiston avulla työn hallittavuus-vaativuus -kyselyn rakennevaliditeettiä ja ajallista mittausinvarianssia tutkitaan rakenneyhtälömallinnuksella.</p> <p>Tulokset viittaavat pääsääntöisesti hyvään rakennevaliditeettiin. Joitain huomioita kuitenkin käsitellään diskussiossa. Lisäksi ajallinen mittausinvarianssi tämän tutkimuksen tulosten perusteella pitää. Tulokset tuovat varmuutta mittarin käyttöön ja täten myös aiempien tutkimuksien tuloksien pohjalta tehtyihin päätöksiin.</p>		
Avainsanat – Nyckelord – Keywords Työstressi, rakenneyhtälömallinnus, rakennevaliditeetti, mittausinvarianssi		
Ohjaaja tai ohjaajat –Handledare – Supervisor or supervisors Jari Lipsanen		
Säilytyspaikka – Förvaringställe – Where deposited		
Muita tietoja – Övriga uppgifter – Additional information Kiitän ohjaajaani yliopiston lehtori Jari Lipsasta, professori Marko Elovainiota ja Eero Mikkilää kommenteista ja neuvoista.		

Sisällysluettelo

1. Johdanto	1
1.1 Työn hallittavuus-vaativuus -malli	1
1.2 Työn hallittavuus-vaativuus -mallin kysely	3
1.2 JCQ:n psykometriset tutkimukset yleisesti	4
1.3 Rajoitteet aiempien tutkimusten otannoissa	5
1.4 Faktorien määrä ja kyselyn tutkimus Suomessa	6
1.5 Tilastolliset menetelmät kyselyn tutkimuksessa	7
1.6 JCQ:n mittausinvarianssi	9
1.7 Tutkimuskysymykset	11
2. Menetelmät	11
2.1 Otos	11
2.2 Arviointimenetelmät	13
2.3 Mittausinvarianssin tutkiminen	15
3. Tulokset	17
3.1 Rakennevaliditeetti.	17
3.2 Mittausinvarianssi	18
4. Pohdinta	23
4.1 Rakennevaliditeetti	23
4.2 Mittausinvarianssi	25
4.3 Rajoitteet	26
4.4 Lopulta	26
5. Lähteet	27
6. Liitteet	34

1. Johdanto

Työstä aiheutuva stressi vaikuttaa ihmisen elämään merkittävästi. Kohonnut työstressi on yhdistetty terveysongelmiin kuten esimerkiksi kohonneeseen verenpaineeseen (McCraty, Atkinson & Tomasino, 2003), sepelvaltimotautien riskiin (Kivimäki ym., 2013) ja kortisolin eritykseen (Dahlgren, Kecklund & Åkerstedt, 2005) ja nämä havaitut yhteydet ovat vain pieni osa jo laajaa tutkimuskenttää. Työstressiä ja uupumusta on tutkittu useissa tutkimuksissa, maissa ja useilla eri ammattiryhmillä esimerkiksi vanginvartijoilla (Schaufeli, & Peeters, 2000), sosiaalityöntekijöillä (Lloyd, King & Chenoweth, 2002) ja sairaanhoitajilla (Lu, Zhao & While, 2019). Siitä on tehty myös Suomessa paljon työstressiin liittyvää tutkimusta (Bordi, Heikkilä-Tammi, Laine, Mäkinen & Seppänen, 2014). Kumuloituva määrä tutkimusta alalla mahdollistaa paremman työelämän ymmärtämisen ja sitä kautta paremman työelämän suunnittelun. Tämä on kriittistä ihmisten työhyvinvoinnin lisäämiseksi tulevaisuudessa. Valtaosa aikuisista ihmisistä käy Suomessa töissä. Mikäli kykenemme tunnistamaan työpaikkojen ominaisuudet, jotka vaikuttavat mielenterveyteen, kykenemme myös tekemään suurissa määrin mielenterveyttä edistävää työtä.

Koska työstressin merkitys yhteiskunnalle/ihmisille on suuri, on sitä syytä kyetä havainnollistamaan käyttökelpoisilla ja selkeillä teorioilla. Työhyvinvointia onkin lähestytty monella tapaa psykologian tutkimuksessa. Aiheesta on kehitetty useita eri teorioita. Esimerkiksi ns. työhön sitoutumisen (work engagement) näkökulma on noussut esille viime aikoina (katso esim. Bakker, Schaufeli, Leiter & Taris, 2008). Tämän näkökulman mukaan henkilöt, jotka ovat ikään kuin työn imussa, ovat tehokkaita ja hyvinvoivia. Toinen huomattavasti vanhempi lähestymistapa on vertailla työn palkitsevuuden ja työpanoksen määrää (Effort-Reward imbalance, ERI; Siegrist, 1996). ERI -mallin mukaan työpanos koostuu ulkoisista vaatimuksista ja henkilön sisäisistä asenteista työtä kohti ja palkitsevuus esimerkiksi rahallisista palkkioista ja statuksesta. Mikäli palkitsevuus on liian pieni suhteessa työpanokseen, on kyseinen työtehtävä henkilölle rasittava.

Tässä tutkimuksessa keskitytään työn hallittavuus-vaativuus -malliin. Sen teoreettinen tausta käydään ensin läpi. Tämän jälkeen tarkastellaan mallin pohjalta tehtyjä psykometrisia tutkimuksia. Lopuksi tutkitaan mallin rakennevaliditeettia ja pitkäikäistä mittausinvarianssia.

1.1. Työn hallittavuus-vaativuus -malli

Työn hallittavuus-vaativuus -malli on yksi tapa kuvastaa työkuormituksen mekanismeja. Mallia on kutsuttu myös työn hallittavuus-vaativuus(-tuki) -malliksi. Tässä tutkimusraportissa käytetään aiempaa. Malli sai alkunsa Yhdysvalloissa ja Ruotsissa 1970-luvulla kerättyjen aineistojen tutkimuksesta (Karasek, 1979). Keskeisenä ajatuksena mallissa on, että eri yhdistelmät työn hallittavuudesta ja vaativuudesta johtavat erilaisiin kokemuksiin työn kuormittavuudesta. Työn hallittavuutta ajatellaan kuvastavan kyky hyödyntää ja kehittää omia taitojaan työpaikalla sekä päättävältä omassa työssä (Karasek, 1985; Karasek ym., 1998). Vaativuutta taas kuvastaa työmäärä, ajalliset paineet ja roolikonfliktit työpaikalla. Eräs tärkeä ominaisuus mallissa on se, että se ottaa huomioon myös työnteon positiiviset vaikutukset eikä pelkästään esimerkiksi työolosuhteista johtuvaa heikentynyttä toimintakykyä.

Karkeasti mallia voidaan kuvata hallittavuuden ja vaativuuden muodostamalla nelijaolla:

- i) alhainen hallittavuus ja vaativuus johtavat *passiiviseen* työhön,
- ii) korkea hallittavuus ja alhainen vaativuus johtavat *alhaiseen stressiin*,
- iii) alhainen hallittavuus ja korkea vaativuus johtavat *korkeaan kuormittavuuteen*
- iv) korkea hallittavuus sekä vaativuus johtavat *aktiiviseen* ja tuottavaan työhön.

Aktiivisessa työssä henkilö on tuottelias produktiivinen ja jatkuvassa kehityksessä. Tämänkaltaisen työ ei ole pitkällä aikavälillä henkilölle terveyden kannalta haitallista. Hallittavuuden laskiessa ja vaativuuksien kasvaessa suhteettomasti henkilö joutuu kuormittavaan työtehtävään ja hänen terveytensä heikkenee. Tätä nelijakoa on kuvattu kuvassa 1. Ideaalitilanteessa henkilö siis tekee aktiivisen työn kategoriaan kuuluvaa työtä. Mallin ajatellaan toimivan universaalisti eri maissa sekä eri ammateissa (Karasek, 1979).

Alkuperäinen malli sisälsi vaativuus- ja hallittavuusdimensiot. Tätä alkuperäistä mallia muokattiin myöhemmin sisältämään kolme uutta dimensiota: sosiaalinen tuki, fyysinen kuormittavuus ja työn epävarmuus (Karasek ym., 1998). Teorian mukaan sosiaalisen tuen, jota saadaan sekä vertaisilta että esimieheltä, määrä vaikuttaa positiivisesti työstä kuormittumisen ehkäisyyn. Samankaltaisesti yleinen fyysinen kuormittavuus ja työn epävarmuus (kuinka suuri todennäköisyys on esim. menettää työpaikka) lisäävät riskiä terveyden heikentymiseen. Työn epävarmuuden dimensiota käytetään harvemmin, koska sitä on vaikea mitata (Karasek ym., 1998). Nykyään ajatellaan olevan siis käytännössä neljä tekijää (vaativuus, hallittavuus, sosiaalinen tuki ja fyysinen kuormittavuus), jotka yhdessä

luovat kokemuksen työn kuormittavuudesta ja ennustavat terveyttä. Lisäksi ensimmäinen dimensio (työn hallittavuus) voidaan jakaa kahtia (päättöksentekoauktoriteetti ja taitojen hyödyntäminen).

Mallia on yleisesti tutkittu huomattavan paljon. Tässä tutkimuksessa keskitytään erään mallin pohjalta luodun kyselyn psykometrisiin ominaisuuksiin. Mikäli lukija haluaa syvemmin paneutua työn hallittavuus-vaativuus -mallin muuhun tutkimukseen, voi suositella esimerkiksi De Langen ja kollegoiden (2003) katsausta mallin pitkäaikaisesta tutkimuksesta, Van der Doefin ja Maesin (1999) katsausta tai uudempaa Häusserin ja kollegoiden (2010) katsausta.

1.2. Työn hallittavuus-vaativuus -mallin kysely

Yksi ensimmäisistä versioista työn hallittavuus-vaativuus -mallin pohjalta luodusta kyselystä julkaistiin vuonna 1985 (Karasek, 1985). Nykyään tämä ”Job Content Questionnaire” (JCQ; Karasek, 1985) on mahdollisesti käytetyin työn hallittavuus-vaativuus -mallin mittari, vaikkakin myös vaihtoehtoisia kyselyjä on olemassa (Karasek, Choi & Ostergren, 2007). JCQ luotiin hyödyntäen aiempien kyselyjen empiiristä ja teoreettista tarkastelua (Karasek ym., 1998). JCQ:sta on myös suomenkielinen versio (Liite B).

Kyselyä on käytetty paljon tutkimuksessa. Esimerkiksi Suomessa JCQ:ta on sovellettu esimerkiksi tutkittaessa aikaisen eläköitymisen ja sairaspöissaolöjen taustoja. Elovainio ja kollegat (2005) käyttivät JCQ:ta aikaisen eläköitymisen harkinnan ennustajana ja havaitsivat työn hallittavuuden ja vaativuuden olevan merkitseviä ennustajia. Heikompi hallittavuus ja korkeat vaatimukset ennustivat aikaista eläköitymisen harkintaa. Taipale ja kollegat (2011) tutkivat sairaspöissaolöjen yhteyttä työn hallittavuuteen ja vaativuuteen. He havaitsivat, että henkilöillä, joilla on korkea työkuormitus JCQ:lla mitattuna (alhainen hallittavuus ja korkeat vaatimukset), on myös kohonnut riski olla enemmän sairaspöissaolöjen takia pöissa töistä/on muita suurempia määriä sairaspöissaolöja.

Mallissa kuvattuja dimensioita mitataan Likert -asteikollisilla (1-5) väittämillä, joita sisällytetään kysymyspatteristoon yhteensä 27 mikäli halutaan hyödyntää lyhyttä versiota kyselystä ja mitata kaikkia dimensioita tai 49 mikäli halutaan hyödyntää pitkää versiota (Karasek, 1985). Psykometrisessä mielessä JCQ:n tulokset muodostavat 2-4 faktoria, joille latautuu kullekin väittämiä vastaavat muuttujat. Mikäli fyysinen kuormitus ja sosiaalinen tuki

otetaan mittaukseen mukaan, muodostuu neljän faktorin malli. Jos mukaan otetaan vain alkuperäiset hallittavuus- ja vaativuusdimensiot, muodostuu kahden faktorin malli. Lisäksi mikäli hallittavuus jaetaan kahtia päätöksentekoauktoriteettiin ja taitojen hyödyntämiseen, faktorien määrä kasvaa yhdellä (alkuperäistä hallittavuusfaktoria ei säilytetä). Lähtökohtaisesti mitään ristiinlatauksia ei tulisi olla havaittavissa, mutta faktorit saavat korreloida keskenään.

1.3. JCQ:n psykometriset tutkimukset yleisesti

Jokaista psykologista mittaria tulee luonnollisesti tarkastella kriittisesti sen pohjalta tehtyjen tutkimusten tulosten valossa. Mittarin validiteetin tulee olla hyväksyttävä mittarin käyttökelpoisuuden takaamiseksi. Seuraavaksi tarkastellaan JCQ:n psykometrisiä ominaisuuksia käsiteltyä tutkimusta. Erityisesti tutustutaan mittarin rakennevaliditeettiin. Rakennevaliditeetti on mittarin ominaisuus, joka kertoo mittaako mittari toivottua asiaa (Cronbach & Meehl, 1955). JCQ:n rakennevaliditeettia on tutkittu useissa eri maissa kuten mm. Kanadassa (Sale & Kerr, 2002), Brasiliassa (Griep ym., 2009; Hökerberg ym., 2010), Ranskassa (Niedhammer, 2002), Koreassa (Eum ym., 2007), Puolassa (Żołnierczyk-Zreda & Bedyńska, 2014) ja Suomessa (Santavirta, 2009; Vastamäki, Vastamäki, Laimi & Saltychev, 2017). Tämä on tärkeää, sillä näin on mahdollista testata hallittavuus-vaativuus -mallin yleistettävyyttä eri maihin. Mikäli malli on yleistettävissä eri populaatioihin, tulisi samankaltainen faktoriratkaisu löytyä kaikista näistä populaatioista.

Yleisesti tulokset viittaavat JCQ:n olevan faktorirakenteeltaan samankaltainen (väittämiä vastaavat muuttujat latautuvat samoille faktoreille) eri maissa, työpaikoissa ja sukupuolten välillä. Kuitenkin joissakin tutkimuksissa faktorirakenne on vaikeasti tulkittavissa (ts. ei vaikuta sopivan teoriaan) suurillakaan otoskoilla (Phakthongsuk & Apakupakul, 2008; Li, Yang, Liu, Xu & Cho, 2004). Esimerkiksi Niedhammer (2002) tutki Ranskassa eksploratiivisella faktorianalyysillä 4 faktorin kyselyn faktorirakennetta. Tuloksissaan hän havaitsi osioiden latautuvan teorian ehdottamille faktoreille (lataukset > 0.4, lukuun ottamatta kahta hallittavuuden osiota), ilman suuria ristiinlatauksia (ristiinlataukset kaikki < 0.2, pääosin < 0.1). Kuitenkin konfirmatorisen faktorianalyysin tulokset ehdottivat mallin sopivan heikosti aineistoon. Lisätutkimusta siis tarvitaan.

1.4. Rajoitteet aiempien tutkimusten otannoissa

Rajoitteena useissa tutkimuksissa on vinoumat otannoissa. Useat tutkimukset on tehty naispainotteisilla otoksilla (esim. Brisson ym., 1998; Li, Yang, Liu, Xu & Cho, 2004; Żołnierczyk-Zreda & Bedyńska, 2014; Eum ym., 2007; Griep ym., 2009; Sale & Kerr, 2002) ja/tai sairaalaolosuhteissa (esim. Sale & Kerr, 2002; Griep ym., 2009; Hökerberg ym., 2010; Eum ym., 2007; Li, Yang, Liu, Xu & Cho 2004). Tämä heikentää tutkimustulosten yleistettävyyttä. Toisaalta myös monissa eri ammateissa työskenteleviä henkilöitä, kuten toimistotyöntekijöitä, kauppiaita ja maanviljelijöitä, sisältäviä tutkimuksia (esim. Żołnierczyk-Zreda & Bedyńska, 2014) ja miehiä kattavasti otoksissaan sisältäviä tutkimuksia (esim. Phakthongsuk & Apakupakul, 2008; Niedhammer, 2002) on tehty.

JCQ:n validiteetin tutkimuksessa olisi suositeltavaa, että aineistoissa on kattavasti eri ammattiryhmistä olevia henkilöitä. Tämä on erityisen tärkeää, koska Karasekin ja kollegoiden (1998) yleisen hallittavuus-vaativuus teorian kuvauksen perusteella tietyt ammattiryhmät sijoittuvat eri kategorioihin rasittavuutensa perusteella (kuormittaviin, ei kuormittaviin, aktiivisiin ja passiivisiin työn laadun kategorioihin). On siis syytä varoa vain yhtä ammattiryhmää sisältävien tutkimuksien tarkastelua tai niistä johtopäätöksien vetämistä, sillä näiden tutkimusten johtopäätökset ovat rajoittuneita kyseessä olevan ammattiryhmän kategoriaan (ts. jos otoksena on sairaanhoitajia, tutkimustulokset yleistyvät ainoastaan aktiivisen laadun kategoriaan (Karasek, 1985)).

Esimerkiksi Sale & Kerr (2002) tutkivat pääasiassa sairaanhoitajia sisältävää aineistoa (n=881, sairaanhoitajien tarkkaa määrää ei ole ilmoitettu) ja havaitsivat tarkasti teorian mukaisen mallin sopivan heikosti aineistoon. Residuaalien välisiä korrelaatioita sallimalla malli sopi tyydyttävästi aineistoon. Salen ja Kerrin (2002) tutkimuksen pohjalta voidaan siis sanoa, että on kohtuullista näyttöä hallittavuus-vaativuus -mallin toimivuudesta mutta vain aktiivisissa työpaikoissa.

Lisäksi on huomioitava, että yhtä ryhmää, kuten sairaanhoitajia, erikseen tutkittaessa on oletettavasti vähemmän relevanttia vaihtelua vastauksissa. Uskottavasti sairaanhoitajilla on hyvin samankaltaiset työn ominaisuudet (esim. hallittavuus ja vaativuus). Näin ollen he vastaavat hyvin samankaltaisesti kyselyyn, joka näitä ominaisuuksia mittaa. Tästä seuraa se, että varianssi joka ilmenee vastauksissa, ei johdu eriävistä työn ominaisuuksista vaan muista tekijöistä (esim. persoonallisuus). Tämä voi mahdollisesti antaa tuloksia tarkasteltaessa vaikutelman kysymysten huonosta erottelukyvystä. Varsinkin jos kaikki vastanneet ovat

samasta toimipisteestä (kuten samasta sairaalasta) on mahdollista, että vaihtelu työolosuhteissa on vähäistä. Yhden ammattiryhmän ja yhden toimipisteen sisällä suoritettun hallittavuus-vaativuus -kyselytutkimuksen tulokset saattavat olla siten harhaanjohtavia.

1.5. Faktorien määrä ja kyselyn tutkimus Suomessa

Työn hallittavuutta mittaavat osiot vaikuttavat latautuvan eri tavoin eri maissa tehdyissä tutkimuksissa. Jotkut tutkimukset ehdottavat hallittavuuden olevan yksi kokonaisuus (esim. Eum ym., 2007; Żołnierczyk-Zreda & Bedyńska, 2014; Brisson ym., 1998; Fujishiro ym., 2011), toiset taas hallittavuuden jakautuvan kahtia päätöksentekoauktoriteettiin ja taitojen hyödyntämiseen (esim. Sale & Kerr, 2002; Hökerberg ym., 2010; Niedhammer, 2002; Kawakami, Kobayashi, Araki, Haratani & Furui, 1995). Esimerkiksi Eum ja kollegat (2007) tutkivat koreankielistä JCQ:ta ja havaitsivat osioiden latautuvan suhteellisen yksiselitteisesti yhdelle hallittavuuden faktorille (lataukset pääosin $> .6$). Toisaalta Hökerberg ja kollegat (2010) havaitsivat kahtiajaetun hallittavuuden faktorin sisältävän mallin sopivan aineistoon, kun taas malli, jossa kahtiajakoa ei oltu tehty, ei sopinut aineistoon.

Suomenkielistä kyselyä on tutkittu kahdesti konfirmatorisilla faktorianalyysillä. Ensimmäisessä tutkimuksessa Santavirta (2009) tutki kyselyä opettajia ($n=1028$) ja sairaanhoitajia ($n=603$) sisältävillä otoksilla. Tutkimuksessaan Santavirta testasi sekä kahden faktorin mallia (ilman hallittavuuden jakamista kahtia päätöksentekoauktoriteettiin ja taitojen hyödyntämiseen) että kolmen faktorin mallia (kahtiajaolla). Mallien yhteensopivuus aineistoon ei ollut yleisesti hyväksyttävällä tasolla. Kolmen faktorin malli vaikutti sopivan aineistoon paremmin, vaikka senkin mallin yhteensopivuus oli heikko.

Santavirran (2009) tutkimuksesta on tehtävä kaksi huomioita: Ensinnäkin tutkimuksessa käytetty pääkomponenttianalyysi ei ole teorian testaamisessa yhtä sopiva menettelytapa kuin faktorianalyysi (Bryant & Yarnold, 1995). Toiseksi aineistoja supistettiin 250 havainnon kokoisiksi ja näiden aineistojen pohjalta tehtiin konfirmatorinen faktorianalyysi. Näin pienet aineistot saattavat aiheuttaa haasteita, kun käytetään konfirmatorista faktorianalyysiä (Tanaka, 1987).

Toisessa Suomessa tehdyssä tutkimuksessa Vastamäki ja kollegat (2017) tutkivat JCQ:n rakennevaliditeettia muusikoita sisältävässä aineistossa ($n=590$). Santavirran (2009) tulosten mukaisesti kahden faktorin malli (ilman hallittavuuden kahtiajakoa) ei sopinut aineistoon.

Myöskään kolmen faktorin malli ei sopinut hyvin aineistoon. Tämän tuloksen pohjalta tutkijat jatkoivat muokkaamalla mallia modifikaatioindeksien perusteella ja lopullinen malli sopi aineistoon kohtuullisesti. Lopullinen malli sisälsi kuitenkin useita residuaalien välisiä korrelaatioita. Tämä tekee tuloksen tulkinnasta haastavaa, sillä lopullinen malli on huomattavasti teorian ehdottamasta mallista poikkeava.

Kokonaaisuudessaan Suomessa tehty tutkimus on heikosti yleistettävissä, sillä se on tehty rajatuilla aineistoilla (opettajilla, muusikoilla ja sairaanhoitajilla). Laaja-alaisempiin aineistoihin perustuvaa tutkimusta tarvitaan, jotta yleistettävissä olevia johtopäätöksiä voidaan tehdä. On siis aiheellista tutkia JCQ:n rakennevaliditeettia lisää Suomessa, sillä sitä on sovellettu jo tutkimuksissa.

1.6. Tilastolliset menetelmät JCQ:n rakennevaliditeetin tutkimuksessa

JCQ:n tutkimuksen suhteen on syytä pitää multinormaalijakautuneisuuden oletus mielessä, sillä kysymykset ovat ordinaalisella mitta-asteikolla. Multinormaalijakautuneisuuden oletus tarkoittaa sitä, että jokaisen muuttujan tulee olla normaalijakautunut, sillä multinormaalisuuden pitäessä kaikki muuttujat ovat normaalijakautuneita (Burdenski, 2000). Tarkalleen ottaen järjestysasteikolliset muuttujat eivät noudata normaalijakaumaa (Micceri, 1989, mutta katso Knapp, 1990), mikä aiheuttaa sen, että oletusta lähtökohtaisesti rikotaan. Kaikki estimointimenetelmät eivät tee tätä oletusta, mutta yleisimmin käytetty suurimman uskottavuuden menetelmä (Bollen, 1989) olettaa muuttujien olevan multinormaalijakautuneita. Vaihtoehtoinen ja myös suhteellisen yleinen menetelmä on ns. DWLS -menetelmä (Diagonally Weighted Least Squares; Muthén, 1984; Muthén, du Toit & Spisic, 1997; Auspharov & Muthén, 2010), joka hyödyntää polykorisia korrelaatioita estimoinnissa. Suurimman uskottavuuden ja DWLS estimointimenetelmien laskennalliset kaavat ovat esitetty liitteessä A.

Mikäli käytetään suurimman uskottavuuden estimointia ja multinormaalijakaumaoletusta on vakavasti rikottu, on lopputuloksena virheelliset estimaatit ja yhteensopivuusindeksit (Mindrila, 2010; Li, 2016). Tämä saattaa johtaa todellisuudessa sopivan mallin hylkäämiseen, sillä yhteensopivuutta kuvastavien indeksien arvot ovat vinoutuneet virheellisten parametriestimaattien johdosta. Edellinen huomioituna on ongelmallista, että suuressa osassa JCQ:n tutkimuksista ei ole raportoitu estimointimenetelmää tai osioiden jakaumien tarkastelua, kun on käytetty suurimman uskottavuuden menetelmää. Lisäksi DWLS -

estimointi vaikuttaa tuottavat korrektit estimaatit ordinaaliasteikollisilla muuttujilla yleisesti paremmin kuin suurimman uskottavuuden menetelmä (Mîndrila, 2010; Li, 2016, mutta katso: Rhemtulla, Brosseau-Liard, Savalei, 2012). Täten JCQ:n rakennevaliditeettia tutkittaessa on perusteltavissa hyödyntää DWLS -estimointia.

Edelliseen päätelmään liittyen JCQ:n tutkimuksessa teorian ehdottaman mallin sopivuus vaikuttaa riippuvan merkittävästi estimointimenetelmästä. Esimerkiksi molemmissa suomalaisissa rakennevaliditeetin tutkimuksissa käytettiin suurimman uskottavuuden menetelmää ja molemmissa tutkimuksissa mallin sopivuus aineistoon oli heikko ennen huomattavia teorian vastaisia muutoksia (Santavirta, 2009; Vastamäki, H., Vastamäki, M., Laimi, & Saltychev, 2017). Lisäksi Sale ja Kerr (2002) sovittivat kyselyn perusteella havaittuun aineistoon teorian mukaisen mallin käyttäen myös suurimman uskottavuuden menetelmää. Mallin yhteensopivuus oli vain kohtalainen yhteensopivuusindeksien perusteella, vaikkakin virhetermien annettiin korreloida teorian vastaisesti eri faktoreihin kuuluvien osioiden välillä ja yksi osio poistettiin kokonaan. Verrattaen Hökerberg ja kollegat (2010) tutkivat teorian ehdottaman mallin sopivuutta samankaltaisesti, mutta käyttivät DWLS -menetelmää. Heidän tulosten perusteella teorian ehdottama malli vaikutti sopivan aineistoon kohtalaisen hyvin ilman suuria modifikaatioita. Vastaavasti Chungkham ja kollegat (2013) havaitsivat mallin sopivan hyvin aineistoon käytettäessä DWLS -menetelmää.

Yhteensopivuusindeksejä tulkittaessa on huomioitava, että eri yhteensopivuusindeksit ovat sensitiivisiä eri tyyppisille mallien määrittelyvirheille. Esimerkiksi SRMR on erityisen herkkä tunnistamaan virheitä latenttien muuttujien välisissä kovariansseissa kun taas CFI ja RMSEA ovat herkkiä painokertoimien virheille (Hu & Bentler, 1998). Tämän vuoksi on syytä tarkastella useita indeksejä, kun tehdään päätöksiä mallin yhteensopivuudesta. On siis suositeltavaa, että raportoidaan kattavasti usea yhteensopivuusindeksi, jotta tuloksista vedetyt johtopäätökset ovat vakuuttavasti perusteltavissa. Tästä huolimatta tutkimuksista osassa on jätetty raportoimatta yhteensopivuusindeksit kokonaan tai raportoitu vain yksi (esim. Li, Yang, Liu, Xu & Cho, 2004; Kawakami, Kobayashi, Araki, Haratani & Furui, 1995; Brisson ym., 1998; Santavirta, 2009).

Tiivistettynä nousee esille kolme tärkeää huomioita JCQ:n testaamiseen liittyen: Ensinnäkin mallin validiteettia tutkiessa on hyödynnetty erinäisiä estimointimenetelmiä, jotka vaikuttavat johtavan systemaattisesti erilaisiin tuloksiin. DWLS -estimointi vaikuttaa johtavan parempaan yhteensopivuuteen kuin suurimman uskottavuuden -estimointi. Toiseksi sillä

kyseessä on Likert -asteikollinen (järjestysasteikollinen) mittari on oletusarvoisesti suositeltavaa käyttää DWLS -menetelmää. Kolmanneksi indeksien raportointi on ajoittain vajavaista. Vaikuttaa siis tärkeältä, että tulevaisuudessa mallin tutkimuksessa huomioidaan edellä mainitut asiat, jotta tuloksien tulkinta olisi mahdollisimman selkeää.

1.7. JCQ:n mittausinvarianssi

Keskeisenä osana tätä tutkimusta on JCQ:n mittausinvarianssin testaaminen. Mittausinvarianssin peruskysymyksenasettelu on: Toimiiko annettu mittari samankaltaisesti eri ryhmien välillä vai onko ryhmien välillä eroa mittarin toiminnassa?

JCQ:ta on hyödynnetty paljon tutkimuksessa (esim. Van der Doef & Maes, 1999; De Lange, Taris, Kompier, Houtman, & Bongers, 2003; Häusser, Mojzisch & Schulz-Hardt, 2010). Sen mittausinvarianssia on kuitenkin tutkittu hyvin vähän.

Mittausinvarianssin tutkimuksen vähäisyys on haitallista, kun tehdään johtopäätöksiä mittarin perusteella. Jos annettu mittari toimii eri tavoin ryhmien välillä, on mahdotonta tulkita mittarilla havaittuja ryhmien välisiä eroja (Vanderberg & Lance, 2000). Esimerkiksi Li ja kollegat (2006) tutkivat sukupuolieroja työn hallittavuuden ja vaativuuden suhteen. Tämänkaltaisessa tutkimuksessa oletetaan, että JCQ toimii samankaltaisesti sekä miehillä että naisilla. Kuitenkin on mahdollista, että mittari on esimerkiksi eri skaalalla miehillä ja naisilla: toinen ryhmä sisältää enemmän vaihtelua kuin toinen. Toisin sanoen toisen ryhmän sisällä näennäisesti pienet erot ovat merkityksellisiä poikkeamia, kun taas toisen ryhmän sisällä suurempikin absoluuttinen poikkeama keskiarvosta voi olla keskihajontoina ilmaistuna pientä. Mittausinvarianssin tutkimus toisi JCQ:n suhteen valoa tämänkaltaisiin mahdollisiin eroihin ryhmien mittauksessa. Vastaavatko esimiehet kyselyyn samankaltaisesti kuin alaiset? Entä esimerkiksi toimistotyöntekijät ja rakennustyömaalla työskentelevät?

Tässä tutkimuksessa tutkitaan ajallista invarianssia. Ajallisessa mittausinvarianssissa ryhmittelevänä tekijänä toimii aikapiste, jolloin kyselyyn on vastattu (Liu ym., 2017). Toisin sanoen testataan toimiiko kyseinen mittari samankaltaisesti myös mittauskertojen välillä. Koska ajallista mittausinvarianssia ei ole tutkittu kattavasti, on pitkittäistutkimusten perusteella tehtyjen johtopäätösten merkitys epäselvä. Jos esimerkiksi on mitattu kahdessa aikapisteessä JCQ:lla koehenkilöiden työtehtävien laatua on oletuksena, että JCQ mittaa samaa asiaa molempina aikapisteinä. Tämän oletuksen rike kyseenalaistaa mittarin validiteetin ja samalla

sen perusteella tehtyjen johtopäätösten paikkansapitävyyden. Jos ajan myötä mittari käyttäytyy eri tavoin, ei voi olla varma siitä, että muutaman vuoden päästä tehdyssä tutkimuksessa mittari vielä toimii.

Ajatellaan hypoteettinen tilanne, jossa on käytetty mittaria A kahdessa aikapisteessä. Sanotaan että aikapisteessä 1 mittari toimii esitetyn teorian mukaisesti, mutta aikapisteessä 2 mittarin rakenne onkin huomattavasti muuttunut (osiot latautuvat eri tavoin faktoreille). Mitä tulkintoja voimme tehdä ilmiön, jota pyrimme tavoittamaan mittarilla A, mittaamisesta? Ehkä mittaustapa on vanhentunut ja vastaajat tulkitsevat kysymykset virheellisesti. Entä jos osiot eivät enää lataudukaan vahvasti annetuille faktoreille? Ehkä työelämä on muuttunut siten, että esimerkiksi kokemus hallittavuudesta on kasvanut siihen pisteeseen asti, että mittari ei ole enää riittävän sensitiivinen erotellakseen hallittavuuden kokemuksia. Epäilemättä työn hallittavuus-vaativuus paradigman mittaaminen on riippuvainen työelämän muutoksista. Jos samaa mittaria käytetään vuosikymmeniä ja havaitaan sen kriteerivaliditeetin heikentyvän (ts. ei korreloi työntekijöiden terveyteen), onko oikein tulkita se teorian ongelmaksi vai mittauksen validiteetin ongelmaksi? Mittausinvarianssin tutkimus antaa näihin kysymyksiin tarvittavaa selkeyttä: Jos invarianssin oletus on voimassa, ei ole ehkä kyse mittaamisen haasteista. Mikäli toisaalta ajallinen invarianssi ei olekaan voimassa, on harkittava mittarin muokkaamista.

Allekirjoittaneen tietämyksen mukaan vain kolme mittausinvarianssin tutkimusta on tehty JCQ:lle. Fujishiro ja kollegat (2011) tutkivat kyselyn mittausinvarianssia työssä olevien Yhdysvalloissa syntyneiden henkilöiden ja maahanmuuttajien välillä (n=3 126). He suorittivat analyysit erikseen hallittavuuden ja vaativuuden skaaloille. Tuloksissaan he havaitsivat hallittavuuden skaalan olevan invariantti maahanmuuttajien ja synnynnäisten yhdysvaltalaisien välillä, mutta vaativuus skaalan olevan ei-invariantti. Kun kahden muuttujan painokertoimet vapautettiin (toisin sanoen tutkittiin osittaista mittausinvarianssia), myös vaativuus skaalan havaittiin olevan invariantti ryhmien välillä. Valitettavasti hallittavuuden kahtiajaosta ei ole mainintaa, joten jää epäselväksi käytettiinkö yhden vai kahden faktorin mallia. Lisäksi tutkimuksessa tyydyttiin pelkästään yksinkertaisimman mittausinvarianssin tason tarkasteluun (mittausinvarianssin tasoista lisää menetelmissä). Chungkham ja kollegat (2013) tutkivat työn hallittavuus-vaativuus –kyselyn pitkästä mittausinvarianssia ja rakennevaliditeettia Ruotsissa. Tutkimuksessa kysely oli mitattu kolmena eri ajankohtana (n=4 913). Aineisto oli laaja-alaisesti kerätty työssäkäyviltä henkilöiltä. Tutkijat havaitsivat eksploratiivisen faktorianalyysin kautta, että hallittavuuden

skaala jakautuu aineistossa kahtia. Konfirmatorinen faktorianalyysi vuorostaan osoitti mallin sopivan aineistoon. Tutkiessaan mittausinvarianssia he havaitsivat, että mallin ajallinen mittausinvarianssi pitää. Lisäksi yksi Item Response Theory (IRT) -tyyppinen mittausinvarianssitutkimus on tehty JCQ:lle (Choi ym., 2009). Tutkimuksessa vertailtiin eri kielisiä kyselyjä ja niiden mittausinvarianssia. Tuloksissa havaittiin, että osiot toimivat eri tavoin eri kielisissä kyselyissä. Esimerkiksi osa eri kielisistä väittämistä olivat keskiarvoisilta vastauksiltaan eriäviä toisistaan.

1.8. Tutkimuskysymykset

Tässä tutkimuksessa tullaan tutkimaan JCQ:n suomenkielistä versiota. Tutkimuksessa tullaan käsittelemään hallittavuus ja vaativuus dimensioita ja niiden rakennevaliditeettia sekä pitkittäistä mittausinvarianssia. Sosiaalisen tuen, fyysisen kuormittavuuden tai työn epävarmuuden dimensioiden väittämiä ei ollut saatavilla. Tutkimuskysymykset ovat:

1. Onko JCQ:n rakennevaliditeetti Suomessa hyväksyttävällä tasolla?
2. Jakautuuko suomenkielisen kyselyn hallittavuusosio kahtia päätöksentekoauctoriteettiin ja taitojen hyödyntämiseen?
3. Onko suomenkielisen kyselyn ajallisen mittausinvarianssin oletus voimassa?

2. Menetelmät

2.1. Otos

Tutkimuksessa käytetty aineisto on suomalaisesta Lasten Sepelvaltimotaudin Riskitekijät -tutkimuksesta (LASERI). Tutkimus alkoi 1978 pilottitutkimuksilla ja ensimmäinen mittava aineiston keruu tehtiin 1980 (Raitakari ym., 2008). Myöhemmin LASERI -tutkimuksen seurantatutkimuksia on tehty useita. Tässä tutkimuksessa keskityttiin erityisesti uusimpien, vuosina 2007 ja 2012 tehtyjen, seurantatutkimusten välisiin JCQ:n rakenteellisiin eroihin/yhtäläisyyksiin sekä mittausinvarianssiin. Alkuperäisestä aineistosta (n= 3596) poistettiin henkilöt, jotka olivat ilmoittaneet olevansa työttömiä, opiskelijoita, pakkolomalla, kotiäitejä/isiä tai muutoin tämän tutkimuksen kohderyhmän ulkopuolella. Aineistoon jäivät säännöllisessä päivätyössä, osa-aikaisessa työssä ja iltatäi yötyössä olevat, epäsäännöllisiä työaikoja tekevät sekä henkilöt, jotka

Taulukko 1. Taustamuuttujien frekvenssit sekä katoanalyysin χ^2 -testisuureet. Testisuureet ilmaisevat oliko 1. ja 2. aikapisteen otoksien taustamuuttujien jakaumissa eroja.

	Naiset (n=503)	Miehet (n=396)	Chi^2
			11.8***
<hr/>			
Asuinympäristö			
Keskusta	81 (16%)	61 (15%)	3.74
Esikaupunki/Lähiö	243 (49%)	191 (48%)	
Maaseutu	171 (35%)	143 (36%)	
<hr/>			
Työaika			
Päivätyö	352 (70%)	274 (69%)	14.26***
Vuorotyö	99 (20%)	61 (15%)	
Epäsäännöllinen	50 (10%)	60 (15%)	
<hr/>			
Vuosittaiset tulot (euroa)			
Alle 20 tuhatta	68 (14%)	17 (4%)	61.46***
Alle 50 tuhatta	386 (79%)	305 (78%)	
Yli 50 tuhatta	33 (7%)	71 (18%)	
<hr/>			
Sääty			
Naimaton	81 (16%)	59 (15%)	4.74
Naimisissa	285 (57%)	237 (60%)	
Avoliitossa	88 (18%)	87 (22%)	
Eronnut/asumuserossa tai leski	47 (9%)	12 (3%)	

* $p < .05$, ** $p < .01$, *** $p < .001$

ilmoittivat työskentelevänsä kotonaan. Osa aineiston havainnoista poistui kadon vuoksi. Analyysieihin päätyneen aineiston koko oli 899 koehenkilöä. Analyysieihin otettiin mukaan ainoastaan ne havainnot, joissa oli kaikki arvot kiinnostuksen kohteena olevista muuttujista. Valitettavasti saatavilla olevasta aineistosta puuttui ikä-muuttuja, joten aineiston ikäjakauman tarkastelu ei ollut mahdollista. Katoanalyysissä verrattiin χ^2 riippumattomuustestillä

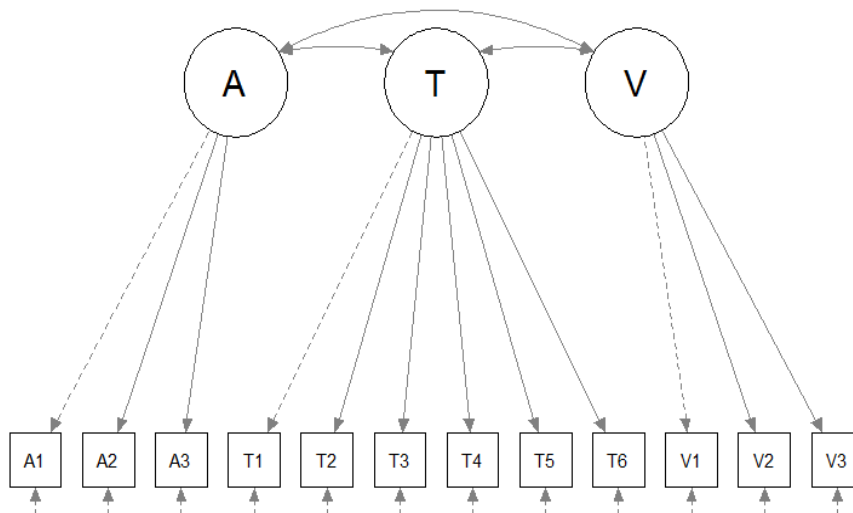
poikkeako tutkimukseen valikoitunut otos ensimmäisen mittauskerran (2007) aineistosta. Testitulokset ovat esitetty taulukossa 1. Testituloksia tarkastellessa vaikuttaa siltä, että analyysiin päätynyt aineisto on valikoitunut systemaattisesti alkuperäisestä aineistosta. Kuitenkin on huomioitavat, että χ^2 testi on erittäin herkkä tunnistamaan eroavaisuuksia suurella otoskoolla.

Aineiston multinormaalisuutta ja yksittäisten osioitten normaalijakautuneisuutta testattiin Henze-Zirklerin (Henze & Zirkler, 1990), Roystonin (Royston, 1983) ja Mardian (Mardia, 1970) multinormaalijakautuneisuuden testeillä, sillä Korkmaz (2014) suosittelee näiden kaikkien testien sisällyttämistä analyysiin. Testit suoritettiin R -ohjelmiston (RC Team, 2013) MVN -paketilla (Korkmaz, 2014). Lisäksi histogrammien tarkastelulla ja Shapiro-Wilk -testeillä tutkittiin yksittäisten muuttujien normaalijakautuneisuutta. Kaikki normalisuuden tai multinormalisuuden testit osoittivat yksiselitteisesti aineiston olevan ei-normaalisti (täten myös ei-multinormaalisti) jakautunut. Lisäksi histogrammien visuaalinen tarkastelu antoi samankaltaisen kuvan: tietyissä muuttujissa oli havaittavissa lievää katto-/lattiaefektiä. Vinouden ja huipukkuuden tunnusluvut olivat 1. aikapisteessä itseisarvoiltaan suurimmillaan 1.08 ja 0.90 sekä 2. aikapisteessä 1.17 ja 1.38 indikoiden lievää poikkeamaa normaalijakaumasta (katso Chou & Bentler, 1995). Aineiston ei-normalisuus ja muuttujien mitta-asteikko (ordinaali) huomioitiin mallinnuksessa sopivan estimointimenetelmän valitsemisella (DWLS).

2.2. Arviointimenetelmät

JCQ:sta (Karasek, 1985; Karasek ym., 1998) käytettiin yhdeksää hallittavuutta ja kolmea vaativuutta kuvastavaa suomenkielistä väittämää. Tutkimuksessa käytettiin siis lyhennettyä versiota kyselystä. Mukana ei ollut fyysisen kuormittavuuden, sosiaalisen tuen eikä työpaikan epävarmuuden osuuksia, sillä niitä ei ollut saatavilla. Kaikki vastaukset väittämiin olivat Likert -asteikolla, vaihdellen 1 ja 5 välillä. Aineiston keruussa koehenkilöitä pyydettiin arvioimaan kuinka hyvin väittämät sopivat heihin (1:”huonosti tai ei ollenkaan”, 2:”melko huonosti”, 3:”tältä väliltä”, 4:”melko hyvin” tai 5:”oikein hyvin”). Tarkat kysymykset ovat esitetty liitteessä B. Osiot ovat nimetty tässä suuraakkosin ja numeroin. Esimerkiksi A1 -osio: ”Voin tehdä paljon itsenäisiä päätöksiä työssäni.”, on päätöksentekoauktoriteettia mittaava väittämä.

2.3. Rakennevaliditeetin tutkiminen



Kuva 1: Määritetty malli yhden aikapisteen sisällä, hallittavuuden kahtiajaolla. Ilman kahtiajakoa, muuttujat A1-A3 sekä T1-T6 latautuivat samalle faktorille. A: päätöksenteko auktoriteetti, T: taitojen hyödyntäminen ja V: vaativuus. Katkoviivalla piirretyt yhteydet ovat lukittu arvoon 1, muut estimoitavia. Kuvassa ei näytetä mittausinvarianssitutkimuksessa käytettyä mallia. Mittausinvarianssitutkimuksessa tämä rakenne toistui kahteen kertaan: Aikapisteen 1 latentit muuttujat regressoituivat vastaaville aikapisteen 2 latenteille muuttujille. Lisäksi residuaalien annettiin korreloida yli ajan vastaavien muuttujien välillä.

Rakennevaliditeettia tutkittiin rakenneyhtälömallinnuksella. Aluksi aikapisteiden sisäisesti tutkittiin faktorirakennetta. Molempien aikapisteiden sisällä sovitettiin erikseen rakenneyhtälömallit, jotka sisälsivät hallittavuuden kahtiajaon, sekä mallit jotka eivät sisältäneet kahtiajakoa. Tämä johti yhteensä neljän eri mallin sovittamiseen: i) 1. aikapisteen sisällä ilman kahtiajakoa ja ii) kahtiajaon kanssa sekä iii) 2. aikapisteen sisällä ilman kahtiajakoa ja iv) kahtiajaon kanssa. Näiden mallien yhteensopivuusindeksit on kuvattu taulukossa 2. Rakenneyhtälömallit määriteltiin Karasekin (1979; 1985) teorian mukaiseksi. Väittämiä vastaavat muuttujat määriteltiin latautuvan asianmukaisille latenteille muuttujille, jotka kuvastivat taitojen hyödyntämistä (T), päätöksenteko auktoriteettia (A) ja vaativuutta (V). Liitteessä B on esitelty muuttujia (joita tullaan nimittämään esim. T1, V2 tai A3) vastaavat väittämät Yhtään ristiinlatausta ei sallittu. Määritettyä rakennetta yhden aikapisteen

sisällä on illustroitu kuvassa 2. Jotta mallit olivat identifioitavissa, lukittiin jokaisen faktorin ja ensimmäisenä sille latautuvan muuttujan välinen yhteys arvoon 1.

Mallien tarkastelussa hyödynnettiin eri yhteensopivuusindeksejä. χ^2 -testisuureen perusteella johtopäätöksiä vetämisestä luovuttiin, sillä se on suurilla otoskoilla hyvin herkkä tunnistamaan mallin yhteensopivuuden virheet ja johtaa käytännössä liian usein mallin hylkäämiseen, vaikka se ei olisi aiheellista (katso esim. Hayduk, Cummings, Boadu, Pazderka-Robinson & Boulianne 2007; mutta myös Barret, 2007). χ^2 -testisuureet ja niiden muutokset on kuitenkin raportoitu. Käytetyt yhteensopivuusindeksit olivat: Tucker Lewis Index (TLI, tunnetaan myös nimellä: Non-Normed Fit Index, NNFI; Tucker & Lewis, 1973), Root Mean Square Error of Approximation (RMSEA; Steiger, 1990) ja Comparative Fit Index (CFI; Bentler, 1990). Hu & Bentler (1999) esittivät kriteerit joiden ylittyessä (tai alittuessa, riippuen tarkasteltavasta indeksistä) malli hylätään kullekin sopivuuden indeksille, kuten $RMSEA > .06$ tai $TLI < .95$. Tulkinnat mallien yhteensopivuudesta tehtiin pitäen nämä kriteerit mielessä, mutta niitä ei käytetty ehdottomina raja-arvoina, kuten on suositeltu (Marsh, Hau & Wen, 2004).

2.4. Mittausinvarianssin tutkiminen

Mittausinvarianssia voidaan testata sekä Item Response Theory (IRT) -tyyppisellä mallinnuksella että rakenneyhtälömallinnuksella (Reise, Widaman & Pugh, 1993). Tässä suoritettiin rakenneyhtälömallinnukseen perustuva mittausinvarianssin testaus.

Jos on olemassa g määrä ryhmiä, joita voidaan mitata jollain mittarilla, jossa on k määrä osioita (havaittuja muuttujia kuten kysymyksiä tai väittämiä), mittausinvarianssissa testattavissa olevat hypoteesit voidaan esittää kaavojen 1 ja 2, jotka kuvastavat yleistä faktorimallia (Jöreskog, 1971), avulla:

$$X_k^g = \tau_k^g + \Lambda_k^g \xi^g + \delta_k^g, \quad (1)$$

Jossa X on mittarin osio, Λ on painokerroinmatriisi, joka liittää latentit muuttujat ξ -osioihin, τ on vakiotermi ja δ virhetermi. Kaavasta 1 voidaan johtaa kaava 2:

$$\Sigma^g = \Lambda_X^g \Phi^g \Lambda_X^{g'} + \Theta_\delta^g, \quad (2)$$

jossa Σ on havaittu kovarianssimatriisi, Φ on latenttien muuttujien välinen kovarianssimatriisi ja Θ on virheiden kovarianssimatriisi. Usein virheiden oletetaan olevan korreloimattomia

keskenään. Tällöin Θ on diagonaalimatriisi, joka sisältää vain virhevarianssit. Asettamalla ryhmien välille yhtäsuuruuden ehtoja voidaan luoda testattavia hypoteeseja. Invarianssin testattavat hypoteesit ovat: Latenttien muuttujien ja havaittujen muuttujien välisten yhteyksien konfiguraalinen samankaltaisuus (havaitut muuttujat ovat yhteydessä samoihin latentteihin muuttujiin jokaisessa ryhmässä, ehtona kaava 3), painokertoimien yhtäsuuruus ryhmien välillä (ehtona kaava 4), vakiotermien yhtäsuuruus ryhmien välillä (ehtona kaava 5), yksilöllisten virhevarianssien yhtäsuuruus ryhmien välillä (ehtona kaava 6) ja latenttien muuttujien kovarianssien yhtäsuuruus (ehtona kaava 7). Vakiotermien yhtäsuuruus on ordinaaliasteikollisella muuttujalla vastaavasti kynnysarvojen yhtäsuuruus (kaava 5, jossa c kuvastaa kynnysarvoa). On myös mahdollista aloittaa mittausinvarianssin testaaminen estimoimalla ryhmien kovarianssimatriisit ja testaamalla niiden yhtäsuuruutta. Mikäli kovarianssimatriisit eivät ole yhtä suuria, tulee jatkaa edellä mainittuja hypoteeseja testaamalla (Vanderberg & Lance, 2000).

Muodostettavissa olevat hypoteesit:

$$\xi^g = \xi^{g'} \quad (1. \text{ taso}) \quad (3)$$

$$\Lambda^g = \Lambda^{g'} \quad (2. \text{ Taso}) \quad (4)$$

$$\tau^g = \tau^{g'}, \text{ tai kynnysarvoilla } \tau_c^g = \tau_c^{g'} \quad (3. \text{ Taso}) \quad (5)$$

$$\theta^g = \theta^{g'} \quad (4. \text{ Taso}) \quad (6)$$

$$\phi^g = \phi^{g'} \quad (7)$$

Näitä hypoteeseja testaamalla voidaan askeltavalla menetelmällä määrittää annetun mittarin mittausinvarianssin taso. Edetessään mittausinvarianssin testaamisessa askeltavasti sisällytetään edellä mainittuja ehtoja analyysiin ja säilytetään vanhat oletukset lisättäessä uusi. Widaman & Reise (1997) esittävät neljä eri mittausinvarianssin tasoa: i) konfiguraalinen, ii) heikko, iii) vahva ja iv) ankara. Näitä tasojen nimityksiä on käytetty paljon tutkimuksessa. Tässä tutkimuksessa tullaan kuitenkin viittaamaan mittausinvarianssin tasoihin yksinkertaisesti 1., 2., 3. ja 4. –tason mittausinvariansseina. 1. Tason mittausinvarianssi viittaa kaavan 3 paikkansa pitävyyteen, 2. tason kaavan 4, 3. tason kaavan 5 ja 4. tason kaavan 6. On huomioitava, että mittausinvarianssia on mahdollista testata myös osittaisesti asettamalla oletuksia muuttujakohtaisesti (Byrne, Shavelson, & Muthén, 1989). Lisäksi voidaan testata faktorien varianssi-kovarianssimatriisien yhtäsuuruutta (kaava 7).

Tätä harvemmin tehdään. Tämän tutkimuksen tulosesiossa tulee esiin se, että faktorien varianssi-kovarianssimatriisien yhtäsuuruutta ei ollut mielekästä testata tässä tutkimuksessa.

Yhteensopivuusindeksien muutoksien tarkastelussa mittausinvarianssin kontekstissa on ongelma, joka tulisi huomioida: Kuinka suuri ero on riittävän suuri, jotta invarianssin oletus hylätään? Joitain viitearvoja on annettu (esim. Chen, 2007) ja niitä tullaan hyödyntämään tässä tutkimuksessa. Lisäksi yhteensopivuusindeksit käyttäytyvät eri tavoin edettäessä invarianssin tasoilla. Esimerkiksi Chen (2007) havaitsi ΔSRMR kasvavan suhteettoman paljon 2. tason invarianssia testattaessa ja Cheung ja Rensvold (2002) havaitsivat 1. tason mallin yhteensopivuuden korreloivan ΔTLI ja ΔRMSEA suuruuden kanssa. Nämä havainnot tekevät SRMR, TLI ja RMSEA -indekseistä hankalasti tulkittavia, kun tarkastellaan eri tasojen mallien välisiä yhteensopivuuseroja.

Myös estimointimenetelmän valinta (suurimman uskottavuuden ja DWLS:n välillä) vaikuttaa mittausinvarianssin testauksessa yhteensopivuusindeksien muutokseen (Sass, Schmitt & Marsh, 2014). Lisäksi Sass ja kollegat (2014) havaitsivat simulaatiotutkimuksessaan, että DWLS -estimointia käytettäessä ja mallin ollessa virheellisesti määritelty (puuttuvia ristiinlatauksia) oli invarianssin pätevyyden tulkitseminen erittäin epävarmaa, sillä yhteensopivuusindeksit heikkenivät itse asiassa vähemmän tilanteessa, jossa ryhmien välinen invarianssi ei pitänyt verrattaen tilanteeseen, jossa invarianssin oletukset pätevät. He suosittelevatkin, että DWLS -estimointia käytettäessä tutkijan ei tulisi nojata tulkinnoissaan liikaa yhteensopivuusindeksien muutoksen tarkasteluun.

Kaikki rakennevaliditeetin/mittausinvarianssin analyysit suoritettiin R-kieleen (RC Team, 2013) muodostetulla Lavaan -paketilla (Rosseel, 2012) hyödyntäen R-studio -käyttöympäristöä (Racine, 2012). Yleisesti rakenneyhtälömallinnusta varten luodut ohjelmat / paketit tuottavat riittävän samankaltaisia tuloksia, jotta niitä voitaisiin verrata keskenään (Narayanan, 2012). Kaikki rakennevaliditeetin ja mittausinvarianssin estimaatit suoritettiin DWLS -menetelmällä.

3. Tulokset

3.1. Rakennevaliditeetti

Kummassakaan aikapisteessä kahden faktorin malli ei sopinut aineistoon ($\text{RMSEA} > 0.1$). Niiden lähempi tarkastelu huonon yhteensopivuuden tähden sivuutetaan täysin. Sen sijaan

kolmen faktorin mallit sopivat hyvin aineistoon. On kuitenkin huomioitava suhteellisen korkeat RMSEA -tasot (0.07 ja 0.08, taulukko 2). Kolmen faktorin malleissa kaikki painokertoimet olivat erittäin merkitseviä ($p < 0.001$). Faktorien väliset kovarianssit olivat myös merkitseviä pois lukien 2. aikapisteen faktorien A ja V välinen korrelaatio ($r = -0.001$, $p = 0.982$). Yksi muuttuja, T2, osoittautui heikosti malliin sopivaksi. Sen residuaalivarianssi oli suuri molemmissa aikapisteissä: 0.802 ensimmäisessä, ja 0.819 toisessa. Vastaavasti selityskertoimet (1-residuaalivarianssi) olivat 0.198 ja 0.181. Modifikaatioindeksit ehdottivat 1. aikapisteen kolmen faktorin mallin yhteensopivuuden paranevan, mikäli ristiinlataus T4 ja auktoriteetti faktorin välillä sallittaisiin (ΔKhi^2 , kun lataus sallittiin = 49.812, $p < 0.001$ ja odotettu standardoitu painokerroin = -0.289). Myös 2. aikapisteessä modifikaatioindeksit osoittivat saman latauksen sallimisen parantavan mallia huomattavasti.

Taulukko 2. Rakennevaliditeetin tutkimuksessa käytettyjen mallien yhteensopivuus

Mallien yhteensopivuus						
Malli	KHI ²	df	p	TLI	CFI	RMSEA
Aika1, 2f	1036.43	53.00	0.00	0.89	0.91	0.14
Aika2, 2f	1175.54	53.00	0.00	0.88	0.90	0.15
Aika1, 3f	272.75	51.00	0.00	0.97	0.98	0.07
Aika2, 3f	369.92	51.00	0.00	0.96	0.97	0.08

Huom. Khi² laskuissa käytetty korjausta (Satorra, 2000).

3.2 Mittausinvarianssi

Mittausinvarianssin tutkimisessa aineistoon sovitettiin kolmen faktorin malli, sillä rakennevaliditeetin tutkimuksessa kahden faktorin mallit eivät sopineet aineistoon. Mittausinvarianssia tutkiessa käytettiin kuvan 1 faktorirakennetta. Erona kuitenkin se, että mittausinvarianssin tutkimuksessa rakenne toistui kahteen kertaan sovitetussa mallissa, sillä molemmat aikapisteet otettiin malliin mukaan samanaikaisesti (katso esim. Widaman, Ferrer & Conger, 2010). 1. Tason malli ei sopinut aineistoon ilman vastaavien osioiden suhteen yli ajan korreloivia residuaaleja ($\text{RMSEA} > .1$). Yli ajan korreloivien residuaalien kanssa, 1.

tason malli sopi aineistoon hyvin (taulukko 4). Residuaalien annettiin siis korreloida vastaavien muuttujien välillä yli ajan (esim. 1. aikapisteessä mitatun muuttujan $T1_{t1}$ yksilöllinen varianssi korreloi 2. aikapisteessä mitatun muuttujan $T1_{t2}$ kanssa). Tämä on perusteltavissa sillä, että samat idiosynkraattiset tekijät vaikuttavat molemmissa aikapisteissä, johtaen systemaattiseen virheeseen osioissa (katso esim. Liu ym., 2017). Sillä virhetermit eivät olleet suuria (taulukko 4.) Tämä ei aiheuta epäselvyyttä tulosten tulkinnassa. Lisäksi 1. aikapisteen latentit muuttujat regressoituvat 2. aikapisteen vastaaville latenteille muuttujille (esim. 1. aikapisteen hallittavuutta kuvastava latentti muuttuja H_{t1} , regressoitui 2. aikapisteen latentille muuttujalle H_{t2}). Aikapisteen sisällä latenttien muuttujien annettiin korreloida keskenään. Mittausinvarianssitutkimusta ei tehty erikseen kahtiajaolla ja ilman, vaan rakennevaliditeetin tutkimuksessa paremmaksi osoittautunut rakenne säilyi mittausinvarianssitutkimukseen. Malli sisälsi molemmat aikapisteet eli yhteensä 6 latenttia muuttujaa. Kaikki havaittujen muuttujien lataukset faktoreille olivat tilastollisesti erittäin merkitseviä ($p < 0.001$). Odotetusti myös mittausinvarianssin 1. tason mallin estimoiduissa parametreissa on havaittavissa selkeästi koholla oleva virhevarienssi T2 muuttujan osalta (aikapisteessä 1: 0.544 ja aikapisteessä 2: 0.557).

Taulukko 3. Mittausinvarianssissa käytettyjen mallien yhteensopivuus, sekä mallien väliset vertailut

Mallien yhteensopivuus						
	KHI ²	df	P	TLI	CFI	RMSEA
1. taso	867.57	231	<0.001	0.969	0.974	0.055
2. taso	862.68	240	<0.001	0.971	0.974	0.054
3. taso	940.40	288	<0.001	0.974	0.973	0.050

Tasojen väliset vertailut						
	Δ KHI ²	df	p	Δ TLI	Δ CFI	Δ RMSEA
1. - 2.	26.287	9	0.002	0.002	>-0.001	0.001
2. - 3.	75.311	48	0.007	0.003	-0.001	0.004
1. - 3.	107.390	57	0.000	0.005	-0.001	0.005

Huom. KHI² laskuissa käytetty korjausta (Satorra, 2000).

Faktorien väliset yhteydet olivat osittain ei merkitseviä 1. tason mallissa. Auktoriteetti ja vaativuus -faktoreiden korrelaatiot eivät olleet merkitseviä kummassakaan aikapisteessä (1.: $r = 0.0723$, $p = 0.090$; 2.: $r = -0.068$, $p = 0.319$). Molemmat kuitenkin jätettiin malliin, sillä niitä voidaan ajatella teorian mukaisina yhteyksinä (Karasek, 1985). Faktorien standardoidut regressiokertoimet vastaaville aikapisteessä 2 kuvastaville faktoreille olivat: A: 0.745, T: 0.765 ja V: 0.687. Faktorien väliset regressiot (ei taulukoitu) olivat kaikki merkitseviä (kaikkien $p < 0.001$). Faktorien varianssit ja niiden keskivirheet (ei taulukoitu) olivat aikapisteessä 1: A: 0.807, 0.022; H: 0.477, 0.028; V: 0.291, 0.042. Vastaavasti aikapisteessä 2: A: 0.377, 0.026;

Taulukko 4. 1. Tason mittausinvarianssin mallin estimoidut lataukset, residuaalivarianssit (e), z -arvot, merkitsevyys sekä standardoidut painokertoimet. Lisäksi, taulukossa on esitetty latenttien muuttujien väliset kovarianssit, niiden keskivirheet, z -arvot, merkitsevyydet sekä korrelaatiot.

Lataukset					
	Estimaatti	Keskivirhe	z	e	Standardoitu painokerroin
Aikapiste 1					
Auktoriteetti					
A1	1			0.134	0.898
A2	1.008	0.021	48.939	0.125	0.906
A3	0.889	0.023	38.373	0.251	0.799
Taidot					
T1	1			0.363	0.691
T2	0.672	0.042	15.894	0.544	0.464
T3	1.096	0.037	29.455	0.297	0.757
T4	1.125	0.039	28.647	0.276	0.777
T5	1.104	0.041	26.902	0.291	0.762
T6	1.195	0.039	30.5	0.222	0.825
Vaativuus					
V1	1			0.492	0.54
V2	0.991	0.098	10.069	0.496	0.535
V3	1.493	0.165	9.074	0.244	0.806
Aikapiste 2					
Auktoriteetti					
A1	1			0.107	0.920
A2	1.01	0.021	48.342	0.095	0.929
A3	0.768	0.024	32.42	0.348	0.707

Taidot					
T1	1			0.371	0.683
T2	0.651	0.048	13.641	0.557	0.445
T3	1.13	0.04	28.132	0.281	0.772
T4	1.155	0.042	27.589	0.263	0.789
T5	1.147	0.044	26.262	0.269	0.783
T6	1.206	0.042	29.009	0.224	0.823
Vaativuus					
V1	1			0.480	0.556
V2	1.043	0.096	10.836	0.460	0.58
V3	1.437	0.151	9.509	0.251	0.799
Faktorien väliset kovarianssit					
	Estimaatti	Keskivirhe	z	P(> z)	Korrelaatio
Aikapiste 1					
A, T	0.423	0.021	20.366	0	0.681
A, V	0.035	0.021	1.693	0.09	0.072
T, V	0.133	0.019	7.101	0	0.357
Aikapiste 2					
A, T	0.178	0.015	11.807	0	0.661
A, V	-0.017	0.017	-0.996	0.319	-0.068
T, V	0.053	0.012	4.343	0	0.296

H: 0.193, 0.017; V: 0.163, 0.028. Kaikki faktorien varianssit olivat tilastollisesti erittäin merkitseviä ($p < 0.001$). On huomioitava, että faktorien varianssit ja kovarianssit aikapisteessä 2 olivat pienempiä faktorien välisistä regressioista johtuen, mikä todettiin sovittamalla malli jossa ei ollut regressiokertoimia. Tämä tekee faktorien varianssien (ja kovarianssien) testaamisesta epämielikästä. Aikapisteiden välillä toisiaan vastaavien osioiden residuaalikorrelaatiot olivat kaikki merkitseviä pois lukien A2 muuttujan korrelaatio ($r = 0.203$, $p = 0.056$). Tämä korrelaatio jätettiin kuitenkin malliin mittausinvarianssin testauksissa, sillä se oli likimain merkitsevä.

Samankaltaisesti kuin rakennevaliditeetin tutkimuksessa, modifikaatioindeksit osoittivat T4 muuttujan ristiinlatauksen sallimisen auktoriteetti faktorille parantavan mallia (esim. korjattu ΔKhi^2 , kun lataus sallittiin aikapisteessä 2: 96.784, $p < 0.001$ ja odotettu standardoitu painokerroin = -0.328). Tämä päti molemmissa aikapisteissä. Myös V3 -muuttujan ristiinlataukset modifikaatioindeksien mukaan olisivat parantaneet mallia jokseenkin (esim. aikapisteessä 2: korjattu $\Delta\text{Khi}^2 = 25.109$, $p < 0.001$ ja odotettu standardoitu painokerroin =

0.201). Standardoidun residuaalikovarianssimatriisin visuaalisessa tarkastelussa mikään muuttuja ei osoittautunut erityisen ongelmalliseksi. Mallit sopivat aineistoon hyvin (taulukko 3), joten modifikaatioindeksien ehdottamia teorian vastaisia muutoksia mallien määrittelyyn ei tehty.

Mittausinvarianssin testaamisessa suurin painoarvo asetettiin ΔCFI :lle. Tämä sen tähden että nykytutkimus on osoittanut sen olevan yksi stabiilimmeistä indekseistä, jota tarkastella mittausinvarianssin kontekstissa (Cheung & Rensvold, 2002; Chen, 2007; Braddy, 2008; Sass, Schmitt & Marsh, 2014). Kuitenkin myös muiden (RMSEA & TLI) yhteensopivuusindeksien muutoksen tarkastelua tehtiin edettäessä invarianssin tasoilla. Eri tasojen mallit sopivat hyvin aineistoon ja suoraan lasketut erotukset yhteensopivuus - indekseissä olivat pieniä (taulukko 3.). Mittausinvarianssin tarkastelua varten luodut mallit, jotka sisälsivät molemmat aikapisteet, itse asiassa sopivat RMSEA:n mukaan paremmin aineistoon kuin rakennevaliditeetin tutkimuksessa käytetyt mallit. Tämä johtuu siitä, että suuremmassa mallissa vapausasteet olivat huomattavasti lukuisammat ($df = 51$ vs. $df = 231$).

Myös ehdotettuihin kriteereihin verrattuna kaikki indeksien muutokset ehdottavat mittausinvarianssin pitävän. ΔTLI oli itse asiassa positiivinen ehdottaen mallin yhteensopivuuden parantumista. Tämä johtui siitä, että TLI huomioi mallin kompleksisuuden vahvemmin suosien enemmän yksinkertaisempaa mallia. $\Delta RMSEA$ tarkastelu ei välttämättä ole suositeltavaa (Meade, Johnson & Braddy, 2008). Kuitenkin, kaikki $\Delta RMSEA$ arvot olivat ehdotettujen kriteerin alapuolella. Erityisesti ΔCFI oli negatiivinen, mutta kuitenkin huomattavan pieni ja alle esimerkiksi Meaden ja kollegoiden (2008) suositteleman $< -.002$

Tason mallia sovitettaessa ratkaisu ei konvergoitunut (estimoitujen parametrien varianssi-kovarianssimatriisi ei ollut positiivi-definiitti) käytettäessä DWLS -estimointia, mutta konvergoitui, kun muuttujia käsiteltiin jatkuvina. Täten konvergoitumisessa ollut ongelma ei todennäköisesti johdu aineiston ja mallin heikosta yhteensopivuudesta vaan estimointimenetelmän laskennallisista operaatioista. 4. tason mittausinvarianssin oletusta ei voida hylätä tämän tutkimuksen perusteella. Tämä viittaa siihen, että estimoitaessa kynnysarvot (vakiotermien sijaan; eli käsiteltäessä muuttujia järjestysasteikollisina) ja lukittaessa aikapisteiden välillä residuaalivarianssit yhtä suuriksi, estimoidut parametrien varianssit olivat virheellisiä. Läheisempi tarkastelu osoittikin residuaalivarianssien varianssien olevan ajoittain negatiivisia, mikä tarkoittaa Heywood -tapauksien ilmentymistä.

4. Pohdinta

Johdannossa läpikäydyissä tutkimuksissa oli havaittavissa metodologisia ongelmia. Monet tutkimukset hyödynsivät suurimman uskottavuuden estimointimenetelmää, mikä ei välttämättä sovellu järjestysasteikollisille muuttujille (Mîndrila, 2010; Li, 2016). Toisissa taas käytettiin DWLS -estimointia. Tämä tekee tulosten vertailusta haastavaa. On kuitenkin todettavissa, että yleisvaltaisesti tutkimukset osoittavat kyselyn rakennevaliditeetin olevan hyväksyttävällä tasolla. Suomalaista tutkimusta on tehty hieman. Huomionarvoisia ovat mahdolliset metodologiset haasteet suomenkielisen kyselyn tutkimuksessa: Estimointimenetelmän valinta, otoksien heikko edustavuus ja useat sallitut residuaalien korrelaatiot mahdollisesti kyseenalaistavat tutkimustuloksien pitävyyden ja yleistettävyyden. Lisäksi työn hallittavuus-vaativuus -mallia on mittausinvarianssin kannalta tutkittu hyvin vähän, vaikkakin mittausinvarianssin oletus on looginen edellytys minkä tahansa mittarin tulosten perusteella tehtyjen päätelmien luotettavuuteen (Vanderbeg & Lance, 2000).

4.1 Rakennevaliditeetti

Ensimmäisenä tutkimuskysymyksenä oli, jakautuuko hallittavuus kahtia suomenkielisessä mittarissa. Tämän rakennevaliditeetin testaamisessa havaittiin ensinnäkin työn hallittavuus-vaativuus -mallin sopivan heikosti aineistoon ilman hallittavuuden kahtiajakoa viitaten siihen että hallittavuus jakautuu kahtia, mikä vastaa ensimmäiseen tutkimuskysymykseen. Tämä löydös on tehty muuallakin eikä se ole sinänsä teorian vastainen, sillä hallittavuuden voidaan ajatella jakautuvan auktoriteetti/taitojen hyödyntäminen osioihin (Karasek, 1985). On kuitenkin huomioitava konteksti, jossa hallittavuus-vaativuus akselit muodostavat yhdessä työn kuormittavuuden ennusteen. Mielenkiintoiseksi kysymykseksi jää: Onko tarvetta muovata teoriaa huomioimaan auktoriteetin ja taitojen hyödyntämisen mahdolliset eriäväisyydet arvioitaessa tietyn henkilön työn kuormittavuutta? Kolmannen akselin lisääminen teoriaan vaikuttaa käytännössä hankalalta. Lisäksi auktoriteetti ja taitojen hyödyntäminen faktorit korreloivat melko vahvasti keskenään ($r > .65$), ja koska osissa tutkimuksista on havaittu kahden faktorin toimivan hyvin, on ehkä syytä säilyttää kahden akselin paradigma. On myös huomioitava, että teoriaan on jo sisällytetty muita ulottuvuuksia (sosiaalinen tuki, fyysiset vaatimukset ja työn epävarmuus), mikä tekee tulkinnasta vähemmän suoraviivaista. Occamin partaveitsi -periaatteen mukaan yksinkertaisempi toimiva

teoria on parempi. Toisaalta akselien lisääminen lisää hienojakoisen informaation määrää, jonka pohjalta voidaan tehdä tarkempaa tutkimusta.

Toiseen tutkimuskysymykseen antaa vastauksen rakennevaliditeetin tutkimus, kun kahtiajako on tehty. Yhteensopivuusindekseistä pääteltynä mittarin rakennevaliditeetti on hyvä. Tämä tarkoittaa sitä, että kaikki tarkastellut indeksit olivat alle tai yli annettujen viitearvojen (Hu & Bentler, 1998). Pääsääntöisesti vakavia ristiinlatauksia ei ollut havaittavissa. Kuitenkin kaksi erityistä ongelmaa oli havaittavissa rakennevaliditeetin tutkimuksessa. Ensinnäkin T2 muuttuja (joka vastaa väittämää "Työhöni kuuluu paljon samanlaisina toistuvia tehtäviä.") vaikuttaa olevan heikosti hallittavuus-vaativuus -malliin sopiva. Sen standardoidut regressiokertoimet olivat pieniä ja virhetermit suuria, mikä viittaisi siihen, että se ei lataudu kunnollisesti taitojen hyödyntämisen faktorille. Modifikaatioindekseissä ei ollut havaittavissa merkittäviä ristiinlatauksia T2 muuttujan suhteen, mikä viittaa siihen, että se ei yksinkertaisesti kuulu kysymyspatteristoon. Tietysti on tulkinnanvaraista, kuinka suuri regressiokerroin on hyväksyttävä, mutta T2 muuttuja oli selityskertoimeltaan huomattavasti muita muuttujia pienempi (keskimäärin n. 0.3 standardoidun yksikön verran pienempi). Myös Chungkham ja kollegat (2013) havaitsivat T2 muuttujan sopivan heikosti faktorirakenteeseen. He itse asiassa tekivät valinnan poistaa tämä muuttuja, mikä paransi mallien yhteensopivuutta. Tämän tutkimuksen tulokset ovat siis konvergenssissa heidän havaintojen kanssa viitaten kokonaisuudessaan siihen, että T2 muuttuja ei sovi kyselyyn eikä kyseessä ole pelkästään suomen kieleen liittyvä seikka. T2 muuttujaa vastaavan väittämän poistaminen tai rekonstruointi olisi suositeltavaa.

Toiseksi T4 muuttuja modifikaatioindeksien mukaan olisi ristiinlatautunut vaativuus -faktorille. Ristiinlatauksen estimoitu standardoitu koko olisi ollut kohtuullisen suuri kaikissa analyyseissä, viitaten mallin olleen lievästi virheellisesti määriteltä. Tämä ei ole yllättävää, kun tarkastellaan semanttisesta näkökulmasta kyseistä muuttujaa vastaavaa väittämää: "Työni vaatii, että opin uusia asioita". Selvästi sekä työn vaatimukset että taitojen kehittäminen esiintyvät väittämässä. Tämän tutkimuksen tulosten perusteella olisi suositeltavaa muokata tätä väittämää esimerkiksi sanavalintoja muuttamalla siten, että se vastaisi vain taitojen kehittämistä.

Aiemmin tässä tutkimusraportissa kritisoitiin suomalaisten tutkimusten metodiikkaa, erityisesti virhevarianssien korrelaatioiden sallimista. Esimerkiksi Santavirta (2009) salli huomattavan määrän virheiden korrelaatioita, jotta mallin yhteensopivuus paranisi. Eräs

selitys tälle olisi, että virheet korreloivat esim. sen tähden, että kysymykset ovat semanttisesti samankaltaisia, eikä niinkään mallin paikkansapitämättömyyden tähden. Tämä ei kuitenkaan ole uskottava selitys virheiden korrelaatioille, kun huomioidaan tämän tutkimuksen tulokset: Virheiden korrelaatioita ei tarvinnut sallia hyvin yhteensopivien mallien saavuttamiseksi. Ongelmana on, että kahdessa suomenkielisellä kyselyllä tehdyssä tutkimuksessa on käytetty eri estimointimenetelmää kuin tässä tutkimuksessa. Sekä Santavirta (2009) että Vastamäki ja kollegat (2017) käyttivät suurimman uskottavuuden menetelmää, kun taas tässä tutkimuksessa käytettiin DWLS -menetelmää. Tämä johtaa siihen, että aiempien suomenkielisten tutkimusten tuloksien suora vertailu tämän tutkimuksen tuloksiin on haastavaa.

4.2 Mittausinvarianssi

Viimeisenä tutkimuskysymyksenä oli, että pitääkö ajallisen mittausinvarianssin oletus? Tehdyistä tutkimuksista Ruotsissa tehty ajallisen mittausinvarianssin tutkimus (Chungkham, Ingre, Karasek, Westerlund & Theorell, 2013) on lähin vertailukohde tätä tutkimusta ajatellen. Käsillä olevassa tutkimuksessa havaittiin yhtenevästi Chungkhamin ja kollegoiden (2013) tutkimuksen tulosten kanssa hallittavuus-vaativuus -mallin olevan invariantti ajan suhteen. Ensimmäisen tason malli sopi aineistoon hyvin. Toisen ja kolmannen tason malleihin siirryttäessä yhteensopivuusindeksit eivät muuttuneet merkittävästi. Chungkham ja kollegat (2013) hyödynsivät Meaden ja kollegoiden (2008) esittämää 0.002 ΔCFI -kynnysarvoa, jonka mukaan tehdä johtopäätös mittausinvarianssin oletuksen voimassaolosta. Mikäli hyödynnämme samaa kynnysarvoa, tämän tutkimuksen tuloksista voimme tulkita mittausinvarianssin oletusten olevan voimassa (kolmanteen tasoon asti), sillä kaikki CFI muutokset ovat pienempiä kuin 0.002.

Kuten aiemmin on mainittu, virheelliset määritykset mallissa aiheuttavat DWLS -estimointia käytettäessä vinoumaa yhteensopivuusindeksien muutoksessa (Sass, Schmitt & Marsh, 2014). Käsillä olevan tutkimuksen suhteen tämä on tärkeä huomio, sillä T4 muuttuja modifikaatioindeksien mukaan olisi latautunut myös vaativuus faktorille. Kysymykseksi jää havaittujen yhteensopivuusindeksien muutoksen totuudenmukaisuus. Toisaalta mallien väliset erot indekseissä olivat erittäin pieniä, mikä viittaa siihen, että on silti turvallista todeta tulosten pohjalta kolmannen tason mittausinvarianssin oletuksen pitävän.

Neljannen tason malli ei konvergoitunut. Viitteitä sen puolesta, että mallin konvergoitumisen ongelma johtuu estimointimenetelmään liittyvistä laskennallisista ongelmista, on esitetty tuloksissa. Ehkä vahvin peruste tämän puolesta on, että käsiteltäessä muuttujia jatkuvina malli konvergoitui (DWLS -estimoinnilla laskien vakiotermit kynnsarvojen sijaan). Johtopäätöksiä vetäminen neljännen tason mittausinvarianssin suhteen tämän tutkimuksen perusteella on haastavaa. Vaihtoehtoisesti olisi mahdollisesti pitänyt suorittaa kaikki mittausinvarianssin analyysit esimerkiksi suurimman uskottavuuden menetelmällä käyttäen ei-normaalisuuden korjauksia. Tämä todennäköisesti olisi kuitenkin johtanut heikompaan yhteensopivuuteen muissa malleissa (Mîndrila, 2010; Li, 2016). On kuitenkin huomioitava, että Chungkham ja kollegat (2013) kykenivät estimoimaan neljännen tason mallin käyttäen DWLS -estimointia. Tästä johtuen ei tutkimusmenetelmien suunnittelun aikana ollut syytä olettaa tämänkaltaista konvergoitumisen ongelmaa. Konvergoitumisen ongelma luokin merkittävän rajoitteen tämän tutkimuksen tuloksiin.

4.3 Rajoitteet

Merkittävänä puutteena tässä tutkimuksessa neljännen tason mittausinvarianssin malli ei konvergoitunut. Tämä todennäköisesti johtuu estimointiin liittyvistä laskennallisista syistä eikä niinkään neljännen tason mallin heikosta yhteensopivuudesta. Neljännen tason mittausinvarianssin oletuksen pitävyys jää kuitenkin kysymykseksi. Lisäksi aineistossa oli saatavilla vain hallittavuus ja vaativuus dimensioiden väittämät. Sosiaalisen tuen ja fyysisen kuormittavuuden dimensioiden tutkimusta ei kyetty tekemään. On havaittu sosiaalisen tuen dimension rakennevaliditeetin olevan heikko (Höckerberg ym., 2010). Sosiaalisen tuen osalta psykometristen ominaisuuksien tutkiminen olisi siis suositeltavaa myös suomenkielisessä kyselyssä.

Mittausinvarianssin tutkimuksen suhteen on vielä useita testattavia hypoteeseja jäljellä. Esimerkiksi onko eri ikäkohorttien välinen mittausinvarianssi voimassa suomenkielisessä kyselyssä? Entä sukupuolten välinen? Näitä hypoteeseja tulisi jatkossa testata, jotta voimme paremmin ymmärtää JCQ:n suomenkielisen version tuloksia.

4.4 Lopulta

JCQ on paljon käytetty kysely. Tästä syystä on aiheellista tutkia sen psykometrisia ominaisuuksia. Tämän tutkimuksen tulokset antavat luotettavuutta JCQ:n käyttöön tulevaisuudessa ja aiempien tutkimusten johtopäätösten pitävyyteen. Mittarin rakennevaliditeetti ja ajallinen mittausinvarianssi vaikuttavat luotettavilta tämän tutkimuksen tulosten perusteella. Erityisenä vahvuutena tutkimuksessa on laaja-alainen, useita ammattiryhmiä edustava otos. Tämä tekee tuloksista hyvin yleistettäviä Suomessa. Lisäksi tutkimus tuo lisävaloa yleisesti JCQ:n rakennevaliditeettiin ja ajalliseen mittausinvarianssiin. Erityisesti jälkimmäinen on arvokas havainto, sillä mittausinvarianssia yleisesti on tutkittu hyvin vähän kyseisen mittarin suhteen.

5. Lähteet

- Asparouhov, T., & Muthén, B. (2010). Simple second order chi-square correction. *Mplus technical appendix*.
- Bakker, A. B., Schaufeli, W. B., Leiter, M. P., & Taris, T. W. (2008). Work engagement: An emerging concept in occupational health psychology. *Work & stress*, 22(3), 187-200.
- Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual differences*, 42(5), 815-824.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, 107(2), 238.
- Bentler, P. M. (1995). *EQS structural equations program manual* (Vol. 6). Encino, CA: Multivariate software, 129.
- Bollen, K. A. (1989). *Structural equations with latent variables* Wiley. New York.
- Bordi, L., Heikkilä-Tammi, K., Laine, N., Mäkinen, J. P., & Seppänen, S. (2014). Psykososiaalisiin kuormitus- ja voimavaratekijöihin liittyvä työhyvinvointitutkimus Suomessa 2010-2013.
- Brisson, C., Blanchette, C., Guimont, C., Dion, G., Moisan, J., Vézina, M., ... & Mäse, L. (1998). Reliability and validity of the French version of the 18-item Karasek Job Content Questionnaire. *Work & Stress*, 12(4), 322-336.

- Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37(1), 62-83.
- Burdenski Jr, T. K. (2000). Evaluating Univariate, Bivariate, and Multivariate Normality Using Graphical Procedures.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464-504.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling*, 9(2), 233-255.
- Choi, B., Bjorner, J. B., Ostergren, P. O., Clays, E., Houtman, I., Punnett, L., ... & Karasek, R. (2009). Cross-language differential item functioning of the job content questionnaire among European countries: the JACE study. *International Journal of Behavioral Medicine*, 16(2), 136-147.
- Chou, C. P., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling.
- Chungkham, H. S., Ingre, M., Karasek, R., Westerlund, H., & Theorell, T. (2013). Factor structure and longitudinal measurement invariance of the demand control support model: an evidence from the Swedish Longitudinal Occupational Survey of Health (SLOSH). *PLoS One*, 8(8), e70541.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, 52(4), 281.
- Dahlgren, A., Kecklund, G., & Åkerstedt, T. (2005). Different levels of work-related stress and the effects on sleep, fatigue and cortisol. *Scandinavian Journal of Work, Environment and Health*, 31(4), 277-285.
- De Lange, A. H., Taris, T. W., Kompier, M. A., Houtman, I. L., & Bongers, P. M. (2003). "The very best of the millennium": longitudinal research and the demand-control-(support) model. *Journal of occupational health psychology*, 8(4), 282.
- Elovainio, M., Forma, P., Kivimäki, M., Sinervo, T., Sutinen, R., & Laine, M. (2005). Job demands and job control as correlates of early retirement thoughts in Finnish social and health care employees. *Work & Stress*, 19(1), 84-92.

- Eum, K. D., Li, J., Jhun, H. J., Park, J. T., Tak, S. W., Karasek, R., & Cho, S. I. (2007). Psychometric properties of the Korean version of the job content questionnaire: data from health care workers. *International archives of occupational and environmental health*, 80(6), 497-504.
- Fujishiro, K., Landsbergis, P. A., Diez-Roux, A. V., Stukovsky, K. H., Shrager, S., & Baron, S. (2011). Factorial invariance, scale reliability, and construct validity of the job control and job demands scales for immigrant workers: the multi-ethnic study of atherosclerosis. *Journal of immigrant and minority health*, 13(3), 533-540.
- Griep, R. H., Rotenberg, L., Vasconcellos, A. G. G., Landsbergis, P., Comaru, C. M., & Alves, M. G. M. (2009). The psychometric properties of demand-control and effort-reward imbalance scales among Brazilian nurses. *International archives of occupational and environmental health*, 82(10), 1163-1172.
- Hayduk, L., Cummings, G., Boadu, K., Pazderka-Robinson, H., & Boulianne, S. (2007). Testing! testing! one, two, three—Testing the theory in structural equation models!. *Personality and Individual Differences*, 42(5), 841-850.
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods*, 3(4), 424.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55.
- Häusser, J. A., Mojzisch, A., Niesel, M., & Schulz-Hardt, S. (2010). Ten years on: A review of recent research on the Job Demand–Control (-Support) model and psychological well-being. *Work & Stress*, 24(1), 1-35.
- Höckerberg, Y. H. M., Aguiar, O. B., Reichenheim, M., Faerstein, E., Valente, J. G., de Jesus Fonseca, M., & Passos, S. R. L. (2010). Dimensional structure of the demand control support questionnaire: a Brazilian context. *International archives of occupational and environmental health*, 83(4), 407-416.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36(4), 409-426.

Karasek Jr, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative science quarterly*, 285-308.

Karasek, R. A., Pieper, C. F., & Schwartz, J. E. (1985). Job content questionnaire and user's guide. Revision 1.1. *Los Angeles: USCLA*.

Karasek, R., Brisson, C., Kawakami, N., Houtman, I., Bongers, P., & Amick, B. (1998). The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. *Journal of occupational health psychology*, 3(4), 322.

Karasek, R., Choi, B., Ostergren, P. O., Ferrario, M., & De Smet, P. (2007). Testing two methods to create comparable scale scores between the Job Content Questionnaire (JCQ) and JCQ-like questionnaires in the European JACE Study. *International journal of behavioral medicine*, 14(4), 189-201.

Kawakami, N., Kobayashi, F., Araki, S., Haratani, T., & Furui, H. (1995). Assessment of job stress dimensions based on the job demands-control model of employees of telecommunication and electric power companies in Japan: reliability and validity of the Japanese version of the Job Content Questionnaire. *International journal of behavioral medicine*, 2(4), 358-375.

Kivimäki, M., Nyberg, S. T., Fransson, E. I., Heikkilä, K., Alfredsson, L., Casini, A., ... & Goldberg, M. (2013). Associations of job strain and lifestyle risk factors with risk of coronary artery disease: a meta-analysis of individual participant data. *Cmaj*, 185(9), 763-769.

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *The R Journal*, 6(2), 151-162.

Li, J., Yang, W., Liu, P., XU, Z., & CHO, S. I. (2004). Psychometric evaluation of the Chinese (mainland) version of Job Content Questionnaire: a study in university hospitals. *Industrial health*, 42(2), 260-267.

Li, C. H. (2016). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological methods*, 21(3), 369.

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior research methods*, 48(3), 936-949.

- Liu, Y., Millsap, R. E., West, S. G., Tein, J. Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological methods*, 22(3), 486.
- Lu, H., Zhao, Y., & While, A. (2019). Job satisfaction among hospital nurses: A literature review. *International journal of nursing studies*.
- Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural equation modeling*, 11(3), 320-341.
- McCraty, R., Atkinson, M., & Tomasino, D. (2003). Impact of a workplace stress reduction program on blood pressure and emotional health in hypertensive employees. *The Journal of Alternative & Complementary Medicine*, 9(3), 355-369.
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of applied psychology*, 93(3), 568.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological bulletin*, 105(1), 156.
- Mîndrila, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society*, 1(1), 60-66.
- Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49(1), 115-132.
- Muthén, B. O. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. *Psychometrika*.
- Narayanan, A. (2012). A review of eight software packages for structural equation modeling. *The American Statistician*, 66(2), 129-138.
- Niedhammer, I. (2002). Psychometric properties of the French version of the Karasek Job Content Questionnaire: a study of the scales of decision latitude, psychological demands, social support, and physical demands in the GAZEL cohort. *International archives of occupational and environmental health*, 75(3), 129-144.

- Phakthongsuk, P., & Apakupakul, N. (2008). Psychometric properties of the Thai version of the 22-item and 45-item Karasek job content questionnaire. *International journal of occupational medicine and environmental health*, 21(4), 331-344.
- Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. *Psychological bulletin*, 114(3), 552.
- Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological methods*, 17(3), 354.
- Sale, J. E., & Kerr, M. S. (2002). The psychometric properties of Karasek's demand and control scales within a single sector: data from a large teaching hospital. *International archives of occupational and environmental health*, 75(3), 145-152.
- Racine, J. S. (2012). RStudio: a platform-independent IDE for R and Sweave. *Journal of Applied Econometrics*, 27(1), 167-172.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of statistical software*, 48(2), 1-36.
- Santavirta, N. (2003). Construct validity and reliability of the Finnish version of the demand-control questionnaire in two samples of 1028 teachers and 630 nurses. *Educational Psychology*, 23(4), 423-436.
- Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(2), 167-180.
- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In *Innovations in multivariate statistical analysis* (pp. 233-247). Springer, Boston, MA.
- Schaufeli, W. B., & Peeters, M. C. (2000). Job stress and burnout among correctional officers: A literature review. *International Journal of stress management*, 7(1), 19-48.
- Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of occupational health psychology*, 1(1), 27.

- Smith, C. S., Tisak, J., Hahn, S. E., & Schmieder, R. A. (1997). The measurement of job control. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 18(3), 225-237.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate behavioral research*, 25(2), 173-180.
- Taipale, S., Selander, K., Anttila, T., & Nätti, J. (2011). Work engagement in eight European countries: The role of job demands, autonomy, and social support. *International Journal of Sociology and Social Policy*, 31(7/8), 486-504.
- Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child development*, 134-146.
- Team, R. C. (2013). R: A language and environment for statistical computing.
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38(1), 1-10.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods*, 3(1), 4-70.
- Van der Doef, M., & Maes, S. (1999). The job demand-control (-support) model and psychological well-being: a review of 20 years of empirical research. *Work & stress*, 13(2), 87-114.
- Vastamäki, H., Vastamäki, M., Laimi, K., & Saltychev, M. (2017). Confirmatory Factor Analysis of the Finnish Job Content Questionnaire (JCQ) in 590 Professional Musicians. *The international journal of occupational and environmental medicine*, 8(3), 174.
- Żołnierczyk-Zreda, D., & Bedyńska, S. (2014). Psychometric properties of the Polish version of Karasek's Job Content Questionnaire. *International Journal of Occupational Safety and Ergonomics*, 20(4), 583-593.

6. Liitteet

A: Estimointimenetelmien kaavat

Estimointimenetelmistä yleisin on ns. suurimman uskottavuuden -menetelmä, jossa pyritään minimoimaan funktion Z arvo:

$$F_{ML} = \log|\Sigma| - \log|S| + \text{tr}(S\Sigma^{-1}) - r, \quad (Z)$$

jossa Σ on estimoitu kovarianssimatriisi, S on otoksen havaittu kovarianssimatriisi ja r on havaittujen muuttujien määrä.

Toinen yleinen vaihtoehtoinen menetelmä, DWLS (Diagonally Weighted Least Squares), on johdannainen WLS -estimointimenetelmästä, joka on esitetty kaavassa Y:

$$F_{WLS} = [s - \sigma(\theta)]'W^{-1}[s - \sigma(\theta)], \quad (Y)$$

jossa s on havaitun polykorisen korrelaatiomatriisin elementit vektorina (ilman kaksoiskappaleita), $\sigma(\theta)$ on mallin estimoidun polykorisen korrelaatiomatriisin elementit vektorina (myös ilman kaksoiskappaleita) ja asymptoottinen parametriestimaattien välinen kovarianssimatriisi, W , on painokerroinmatriisi. W -matriisin käyttö tarkoittaa käytännössä vinouden ja huipukkuuden, eli normaalijakaumasta poikkeavuuden, huomioimista estimoinnissa. Sen sijaan, että käytettäisiin täyttä painomatriisia W , DWLS -estimoinnissa käytetään osittaista painomatriisia, jossa on ainoastaan diagonaaliset elementit WLS -estimoinnissa käytetystä W matriisista, parametrien estimointiin. Pelkästään diagonaalin hyödyntäminen lieventää laskennallisia vaatimuksia huomattavasti.

B: Suomenkieliset väittämät, joita käytettiin tutkimuksessa

Ensin on esitetty analyyseissä käytetyn muuttujan nimi ja seuraavana sitä vastaava väittämä.

A1: "Voin tehdä paljon itsenäisiä päätöksiä työssäni."

A2: "Minulla on paljon sananvaltaa omiin töihini."

A3: "Minulla on hyvin vähän vapautta päättää, miten teen työni"

T1: "Työni edellyttää minulta luovuutta"

T2: "Työhöni kuuluu paljon samanlaisina toistuvia tehtäviä."

T3: "Työni vaatii pitkälle kehittyneitä taitoja"

T4: "Työni vaatii, että opin uusia asioita"

T5: "Työssäni saan tehdä paljon erilaisia asioita"

T6: "Minulla on mahdollisuus kehittää minulle ominaisia kykyjäni"

V1: "Täytyykö teidän kiirehtiä suoriutuaksenne työstänne?"

V2: "Onko työssänne vaiheita, jolloin työ on liian vaikeaa?"

V3: "Onko työnne henkisesti rasittavaa?"